

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平10-302064

(43)公開日 平成10年(1998)11月13日

(51)Int.Cl.⁶

G 0 6 T 5/20
1/00

識別記号

F I

G 0 6 F 15/68
15/62

4 0 0 A
3 8 0

審査請求 有 請求項の数16 O L (全 17 頁)

(21)出願番号 特願平10-75808

(22)出願日 平成10年(1998) 3月24日

(31)優先権主張番号 60/041695

(32)優先日 1997年3月25日

(33)優先権主張国 米国 (U S)

(31)優先権主張番号 60/043540

(32)優先日 1997年4月14日

(33)優先権主張国 米国 (U S)

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 デイビット ジェイコブス

アメリカ合衆国、 ニュージャージー

08540、 プリンストン、 インディペンデ

ンス ウェイ 4 エヌ・イー・シー・リ

サーチ・インスティテューテュ・インク内

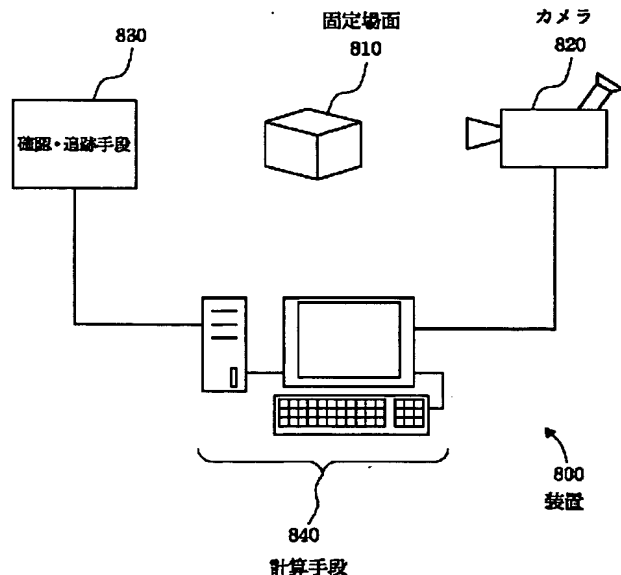
(74)代理人 弁理士 若林 忠 (外4名)

(54)【発明の名称】 欠陥データとの線形適合方法、動きから構造への適用方法、強度画像の特徴決定への適用方法および装置

(57)【要約】

【課題】 欠陥要素を持つ行列に低階数行列を適合させる方法の実現。

【解決手段】 第1段階では、移動カメラを用いて固定場面の画像のシーケンスを得ることで、動きシーケンスから点位置からなる第1の行列を構築し、そのシーケンスから点の素性を確認・追跡し、その素性の座標を用いて、一部の素性が一部の画像に存在しない一部の欠陥要素を有する第1の行列を構築する。第2段階では、第1の行列から列トリプルを選択し、その零空間を第2の行列とし、第2の行列の3つの最小成分を取ることで、近似解を得る。第3段階では、その3つの最小成分に反復アルゴリズムを適用して、第3の行列を構築し、推定を向上させる。最後に第4段階で、第3の行列を分解して、完全な場面構造を求める。



1

【特許請求の範囲】

【請求項1】 不完全なデータを提供するビデオシーケンスから完全な場面構造を形成する方法において、

(a)

(i) 移動するカメラを用いて固定場面の画像シーケンスを得て、

(ii) そのシーケンスから点素性を確認・追跡し、

(iii) 該点素性の座標を用いて、一部の素性が一部の画像で存在しない一部の欠陥要素を有する第1の行列を構築して、ビデオシーケンスから点位置からなる第1の行列を構築する段階と、

(b)

(i) 第1の行列から列トリプルを選択し、

(ii) その零空間から第2の行列を得て、

(iii) 第2の行列の3つの最小の成分を得て、近似解を構築する段階と、

(c) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階と、

(d) 第3の行列を分解して完全な場面構造を求める段階と、を有してなる方法。

【請求項2】 前記第1の行列が下記式で与えられる請求項1記載の方法。

【数1】

$$\hat{Q} \equiv \begin{pmatrix} \hat{X}_{1,1} & \hat{X}_{1,2} & \cdots & \hat{X}_{1,n} \\ \hat{Y}_{1,1} & \hat{Y}_{1,2} & \cdots & \hat{Y}_{1,n} \\ \hat{X}_{2,1} & \hat{X}_{2,2} & \cdots & \hat{X}_{2,n} \\ \hat{Y}_{2,1} & \hat{Y}_{2,2} & \cdots & \hat{Y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \hat{X}_{\frac{m}{2},1} & \hat{X}_{\frac{m}{2},2} & \cdots & \hat{X}_{\frac{m}{2},n} \\ \hat{Y}_{\frac{m}{2},1} & \hat{Y}_{\frac{m}{2},2} & \cdots & \hat{Y}_{\frac{m}{2},n} \end{pmatrix}$$

【式中、

【数2】

$$\hat{Q}$$

は全ての追跡画像点を含む単一の行列であり、

【数3】

$$\hat{Q}$$

の各列には、単一の点の全ての画像座標がある。 $X_{i,j}$ は、 i 番目の画像における j 番目の雑音性像点の X 座標であり、 $Y_{i,j}$ は i 番目の画像における j 番目の画像点の Y 座標である。]

【請求項3】 列トリプルが無作為に選択される請求項1記載の方法。

【請求項4】 第2の行列が下記式で与えられる請求項1記載の方法。

$$N = [N_1 N_2 \cdots N_1]$$

【式中、 S_i は列の i 番目のトリプルによって囲まれる線

2

形空間であり、 N_i は S_i の零空間を表しており、従って N_i の各列は、空間 S_i に対して直交するベクトルを表す。]

【請求項5】 反復アルゴリズムが下記式で与えられる請求項1記載の方法。

【数4】

$$T_{i+1} = \hat{Q} P_i^+$$

【式中、 T_i は、カメラの動きを表す行列 T の反復 i での推定値であり、 P_i は、場面構造を表す行列の反復 i での推定値であり、 P_i^+ は P_i の擬逆行列であり、

【数5】

$$\hat{Q}$$

は全ての追跡画像点を含む単一の行列であり、

【数6】

$$\hat{Q}$$

の各列には、単一の点の全ての画像座標がある。 $X_{i,j}$ は、 i 番目の画像における j 番目の雑音性像点の X 座標であり、 $Y_{i,j}$ は i 番目の画像における j 番目の画像点の Y 座標であって、

【数7】

$$\hat{Q}$$

、 T および P についての一連の推定値が得られ、

【数8】

$$\hat{Q}$$

— Q の大きさは非増加的である。]

【請求項6】 近似解を構築する段階が、

(a) 列の完全に占有された亜矩形の零空間を計算して、列トリプルの零空間を与え、その亜矩形が実質的に縮重している場合に零空間を無視する小段階と、

(b) 何らかの所定の一定因数だけ、第1の行列より零空間が大きくなるまで列トリプルを選択し、第2の行列が所定の閾値より小さい第4の最小特異値を持つ場合には全ての結果を信頼性のないものと判定する小段階と、

(c) データの部分集合について正確な推定値が得られる場合には、反復最小自乗法によって第2の行列を発見的に拡張する小段階と、を有してなる請求項1記載の方法。

【請求項7】 不完全なデータを提供するビデオシーケンスから完全な場面構造を形成する方法において、

(a)

(i) 独立の複数の点光源で、固定場面の一連の画像を得て、

(ii) 付影または反射のために信頼性のないデータを提供する各画像の点を確認し、

(iii) データの信頼性のない欠陥要素を用いて、強度値の第1の行列を構築して、複数の画像から強度値からなる第1の行列を構築する段階と、

(b)

(i) 第1の行列から列トリプルを選択し、

3

- (ii) その零空間から第2の行列を得て、
 (iii) 第2の行列の3つの最小の成分を得て、
 近似解を構築する段階と、

- (c) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階と、
 (d) 第3の行列を分解して完全な場面構造を求める段階と、を有してなる方法。

【請求項8】 第1の行列が下記式で表される請求項7記載の方法。

【数9】

$$\hat{Q} = HN + E$$

【式中、

【数10】

$$\hat{Q}$$

はj番目の列が各画像のj番目の画素の強度を与えるデータ行列であり、Eは感知誤差の行列であり、Hはi番目の行がi番目の画像のhベクトルである照明行列であり、hはその強度によってスケールを決定される点光源の3次元ベクトルであり、Nはj番目の列が $p_j n_j$ である場面行列であって、 p_j は表面アルベドであり、 n_j は3次元面状部分の表面法線である。]

【請求項9】 列トリプルが無作為に選択される請求項7記載の方法。

【請求項10】 第2の行列が下記式で与えられる請求項7記載の方法。

$$N = [N_1 N_2 \dots N_1]$$

【式中、 S_i は列のi番目のトリプルによって囲まれる線形空間であり、 N_i は S_i の零空間を表しており、従って N_i の各列は、空間 S_i に対して直交するベクトルを表す。]

【請求項11】 反復アルゴリズムが下記式で与えられる請求項7記載の方法。

【数11】

$$T_{i+1} = \hat{Q} P_i^+$$

【式中、 T_i は、照明を表す行列Tの反復iでの推定値であり、 P_i は、場面点での表面法線を表す行列の反復iでの推定値であり、 P_i^+ は P_i の擬逆行列であり、

【数12】

$$\hat{Q}$$

は全ての追跡画像点を含む単一の行列であり、

【数13】

$$\hat{Q}$$

の各列には、単一の点の全ての画像座標がある。 $X_{i,j}$ は、i番目の画像におけるj番目の雑音性像点のX座標であり、 $Y_{i,j}$ はi番目の画像におけるj番目の画像点のY座標であって、

【数14】

$$\hat{Q}$$

、TおよびPについての一連の推定値が得られ、

4

【数15】

$$\hat{Q}$$

—Qの大きさは非増加的である。]

【請求項12】 近似解を構築する段階が、

(a) 列の完全に占有された亜矩形の零空間を計算して、列トリプルの零空間を与え、その亜矩形が実質的に縮重している場合に零空間を無視する小段階と、

(b) 何らかの所定の一定因数だけ、第1の行列より零空間が大きくなるまで列トリプルを選択し、第2の行列が所定の閾値より小さい第4の最小特異値を持つ場合には全ての結果を信頼性のないものと判定する小段階と、

(c) データの部分集合について正確な推定値が得られる場合には、反復最小自乗法によって第2の行列を発見的に拡張する小段階と、を有してなる請求項7記載の方法。

【請求項13】 不完全なデータを与えるビデオシーケンスから完全な場面構造を与える装置において、固定場面の画像シーケンスを得るためのカメラと、そのシーケンスから点素性を確認・追跡する手段と、動きシーケンスから点位置からなる第1の行列を構築し、一部の素性が一部の画像で存在しない素性の座標を用いて一部の欠陥要素を有する第1の行列を構築し、

1) 第1の行列から列トリプルを選択する段階、2) その零空間から第2の行列を得る段階、3) 第2の行列の3つの最小の成分を得る段階、4) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5) 第3の行列を分解して完全な場面構造を求める段階を有する近似解の構築を行う計算手段を有する装置。

【請求項14】 計算手段がパーソナルコンピュータである請求項13記載の装置。

【請求項15】 不完全データを提供するビデオシーケンスから完全な場面構造を得る装置において、固定場面の一連の画像を得るための複数の異なる点光源と、ならびに複数の画像からの強度値からなる第1の行列を構築し、付影または反射のために信頼性のないデータを提供する各画像中の点を確認し、データの信頼性がない欠陥データを有する強度値の第1の行列を構築し、

1) 第1の行列から列トリプルを選択する段階、2) その零空間から第2の行列を得る段階、3) 第2の行列の3つの最小の成分を得る段階、4) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5) 第3の行列を分解して完全な場面構造を求める段階を有する近似解の構築を行う計算手段と、を有してなる装置。

【請求項16】 計算手段がパーソナルコンピュータである請求項15記載の装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、欠陥データとの線

形適合、動きから構造への適用、および強度画像の特徴決定への適用に係る方法および装置に関するものである。

【0002】

【従来の技術】いくつかの映像上の問題は、低次元の線形表面のデータへの適合の問題に帰することができる。例えば、トマシら (C. Tomasi et al. (1992) "Shape and Motion From Image Streams Under Orthography: A Factorization Method" Int. J. of Comp. Vis. 9(2):137-154) は、アフィン作像モデルを仮定して、雑音性画像から構築されるデータ行列に階数3行列を適合させることによって、2次元画像のシーケンスで追跡される3次元点集合の構造および動きを求める方法を示している。ベルフメウアらは (Berlhumeur et al. (1996) "What is the Set of Images of an Object Under All Possible Lighting Conditions?" IEEE Conf. of Comp. Vis., and Pat. Rec.:270-277) は、やはり雑音性データから構築される行列に対して階数3行列を適合させることで、凸面ランベルト対象が与える強度画像集合の特徴決定方法を示している。この研究に基づく解法における一つの重大な問題は、一部欠陥要素を有するデータ行列を扱う方法である。動きから構造への場合、ほとんど不可避で起こることであるが、動きシーケンスを通じて画像で一部の素性を見ることができない場合、欠陥要素がデータ行列に生じる。ベルフメウアら (P. Belhumeur et al., 1996) によって提案された方法に従って強度の複写を構築するには、一部の場面点の表面法線が一部の画像で光源に対向しない場合、欠陥行列要素が生じる。それは、平滑な対象を含む場面について起こるはずであり、ほとんどの実際の場面で起こる可能性が高い。

【0003】最近、いくつかの重要な映像上の問題が、それらの問題を、何らかの重要な場面構造を示す低次元の線形表面もしくはアフィン表面検出の問題に帰するアフィン作像モデルを用いることで扱われるようになっていく。例えば、ウルマンら (S. Ullman et al. (1991) "Recognition by Linear Combinations of Models" IEEE Trans. PAMI 13(10):992-1007) およびバスリら (R. Basri et al. (1988) "The Alignment of Objects With Smooth Surfaces" Computer Graphics, Vision and Image Processing: Image Understanding 57(3):331-345) の一次結合では、各種視点からの3次元構造によって得られる画像集合が、全ての可能な画像の空間で低次元アフィン表面と称される。この構造を対象認識に用いる。トマシら (C. Tomasi et al. (1992)) は、これに関連する考え方をを用いて、長い動きシーケンスにより動きから構造を有効に形成している。ノイズのない画像はいずれも低次元線形表面にあるはずであることから、そのような表面をデータに適合させて、それを用いて構造と動きを導き出すことができる。モーゼス (Y. Moses (1993) Face recognition: generalization to novel images Ph.D. The

sis, Weizmann Institute of Science) およびシャシュア (A. Shashua (1992) Geometry and Photometry in 3D Visual Recognition MIT TR-1401) は、単一の視点からであるが、点光源の位置および強度を変えながらの3次元ランベルト構造によって得られる強度画像集合も、可能な画像の空間における3次元線形表面として説明できることを示している。これらは、可能な画像の空間に3次元線形表面を形成する(強度複写)。ベルフメウアらは (P. Belhumeur et al. (1996))、この結果を用いて、凸形状対象が与え得る強度画像の集合について説明し、複数光源と自己付影を可能としている。この研究のいずれにおいても、妥当な仮定集合を見だし、それを線形とすることで、問題はかなり単純化される。

【0004】

【発明が解決しようとする課題】これらの方法はいずれも、データ集合に対して低階数 r の行列を適合させる必要がある。この行列によって囲まれる線形空間は場面構造に相当する。いくつかのアプローチではそれは、ノイズの存在下であっても正確に階数 r を有する $r \times n$ の行列を構成する必要がある最小量のデータのみを獲得することで行われる (S. Ullmanら (1991), Y. Moses (1993), A. Shashua (1992) および P. Belhumeurら (1996) 参照)。この方法は簡単であるが、ノイズに対して非常に敏感であり得る。ウルマンら (Ullman et al. (1991)) も、特異値分解 (SVD) を用いて、追加画像を用いて問題を過度に決定する場合に作られる比較的大きな行列に低階数行列を適合させる方法を提案している。トマシら (C. Tomasi et al. (1992)) は、長い動きシーケンスを用いて、大きな行列を作り、次にSVDを用いて最も近い低階数行列を見いだしている。この方法は、新たな行列における各要素とデータ行列の間の差の2乗の合計を最小とする低階数行列を見いだすものである。問題を過度に決定し、そのデータに対する最小二乗適合を見いだすことで、この方法はノイズの影響を低減することができる。

【0005】これらの方法で重要となる制限は、それらが欠陥要素のないデータ行列を必要とすることである。動きから構造の場合、それは、各点素性があらゆる画像で見ることができなければならないことを意味している。多くの場合、点は動きシーケンス時に視野から出入りするか、あるいはシーケンス途中で他の対象によって視界から遮断されることから、これは厳しい仮定である。ある場面によって得られる強度複写を導く方法は、各画像における照明が単一の点光源から来る画像集合を用いるものである。場面の点が光源によって照明されない場合、そのデータ行列は欠陥要素を持つことになる。従ってこれらの方法では、あらゆる目視可能な場面点をあらゆる画像で照射する必要がある。しかしながら、滑らかな対象がある場面では、その仮定は、光源の方向が目視方向と同じ場合に限って満足される。すなわち、1

個の画像が可能である。より一般的には、光源が目視方向近くにある場合にのみ、ほとんどの可視点が照射される。しかしながら、画像が類似の光源方向で得られた場合は、得られる解は不安定になり得る。

【0006】トマシら (C.Tomasi et al. (1992)) は、欠陥要素を有する行列の簡単な発見的適合方法について説明している。シュームらは (H.Shum et al. (1995) "Principal Component Analysis With Missing Data and Its Applications to Polyhedral Object Modeling" IE EE Trans. PAMI 17(9):854-867) は、ワイバークによる反復法 (T.Wiberg (1976) "Computation of Principal Components When Data Are Missing" Proc.Second. Symp. Computational Statistics:229-236) を関係する動きから構造問題に適用している。この方法は、適合させた低階数行列とデータ行列では欠落していない要素との間の差の2乗の合計を最小とするものである。このアプローチは、局所的に最適解に収束するという利点を持っている。しかしながら、全体的に最適なものが見いだされることを保証するものではない。この反復法には利点があるが、それが良好な開始点を持たない場合は、誤った答えに収束する場合が多い。

【0007】ウルマンら (S.Ullman et al (1991)) およびトマシら (C.Tomasi et al. (1992)) によって提案された方法は、一致した像点のシーケンスが与えられ、それが、アフィンであるかアフィンの部分集合 (例: R. Basri (1996) "Paraperspective Affine" Int.J. of Comp. Vis. 19(2):169-179に記載されているような弱い正射遠近またはparaperspective) であるカメラモデルを用いて3次元から2次元に投影されたものと仮定するものである。多数の点の中心の移動を推定することで、動きの移動成分が別個に計算されることから、我々は単純化のために、面内移動がないと仮定する。 $p_i = (x_i, y_i, z_i)^T$ をi番目の3次元点素性とする。次に、画像変換に対するモデルは 2×3 行列 S_i と書くことができる。 p_i によって与えられる像点を $q_i = (X_i, Y_i)$ で表すと、以下のようになる。

【0008】

【数16】

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

正投影の場合、 S_i は回転行列の最初の2行である。

【0009】そこで我々は、簡単な行列表記で画像シーケンス全体を表わすことができる。i番目の列が p_i である行列を P とすると、以下のようになる。

【0010】

【数17】

$$P \equiv \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ z_1 & z_2 & \cdots & z_n \end{pmatrix}$$

全ての変換 S_i の全ての行を含む行列を T で表すと、 T の $2i-1$ 行が S_i の最初の列であり、 T の $2i$ 行がその次の列である。

【0011】

【数18】

$$T \equiv \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \\ \vdots \end{pmatrix}$$

最後に、全ての追跡画像点を含む単一の行列を

【0012】

【数19】

$$\hat{Q}$$

で表す。

【0013】

【数20】

$$\hat{Q}$$

の各列には、単一の点の全ての画像座標がある。すなわち、

【0014】

【数21】

$$\hat{Q}_{2-i,j}$$

は画像 i における q_j の x 座標を与え、

【0015】

【数22】

$$\hat{Q}_{2-i,j}$$

は画像 i における q_j の y 座標を与える。そして、次のように書ける。

【0016】

【数23】

$$\hat{Q} \equiv \begin{pmatrix} \hat{X}_{1,1} & \hat{X}_{1,2} & \cdots & \hat{X}_{1,n} \\ \hat{Y}_{1,1} & \hat{Y}_{1,2} & \cdots & \hat{Y}_{1,n} \\ \hat{X}_{2,1} & \hat{X}_{2,2} & \cdots & \hat{X}_{2,n} \\ \hat{Y}_{2,1} & \hat{Y}_{2,2} & \cdots & \hat{Y}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \hat{X}_{\frac{m}{2},1} & \hat{X}_{\frac{m}{2},2} & \cdots & \hat{X}_{\frac{m}{2},n} \\ \hat{Y}_{\frac{m}{2},1} & \hat{Y}_{\frac{m}{2},2} & \cdots & \hat{Y}_{\frac{m}{2},n} \end{pmatrix}$$

式中、

【0017】

【数24】

$$(\hat{x}_{i,j})$$

は、i番目の画像におけるj番目の雑音性像点のX座標であり、

【0018】

【数25】

$$\hat{y}_{i,j}$$

も同様に定義される。

【0019】

【数26】

$$\hat{Q}$$

と同じ次元の数列であって、その要素が

【0020】

【数27】

$$\hat{Q}$$

の相当する要素を感知する上で遭遇したノイズを示すものをEとする。そうすると、画像シーケンス全体の形成を次のように簡単に記述することができる。

【0021】

【数28】

$$TP+E=\hat{Q}$$

さらに、次のように定義する。

【0022】

【数29】

$$Q=\hat{Q}-E=TP$$

誤差のないデータQに直接アクセスしないが、

【0023】

【数30】

$$\hat{Q}$$

からそれを推定したい。Tは3つの列（以下列トリプルもしくは端にトリプルという）を持つことから、階数3のみを持つことができ、結果的にQは階数3のみを持つことができる（すなわち、点群と見なされるQの列または行は、 R^M の3次元線形部分空間（ R^D ）以下を占めるに違いない。いずれの列または行も、他の3つの一次独立列もしくは行の一次結合である。）。

【0024】感知される像点の座標を含む行列

【0025】

【数31】

$$\hat{Q}$$

は、Qの雑音性のものであり、最大階数を有する。しかしながら、

【0026】

【数32】

$$\hat{Q}$$

に最も近い階数3のQ行列を見いだせば、Qの最優推定

を行うことができる。詳細には、SVDを用いて、

【0027】

【数33】

$$\sum_i \sum_j (\hat{q}_{ij} - q_{ij})^2$$

を最小とするQ行列を見いだすことができる。さらに、SVDを用いてこのQ行列をTとPに分解して、アフィン変換まで、3次元点の構造および動きの推定を行うことができる（詳細については、C.Tomasi et al. (1992) 参照）。

【0028】一部の画像で一部の点を見ることができない場合、

【0029】

【数34】

$$\hat{Q}$$

はそれらの点についての欠落記述項を含むことになる。さらに、全て見ることができたとしても、誤差のないデータが3階数のQ行列を生じる可能性のあることがわかって（さらに、隣接する画像にある点群の中心の移動を計算することで、データからの隣接する画像間の移動の効果を排除することができる。そのためには、一部の点がある画像から次の画像に常に存在すると仮定するのみでよい。）。従って、現在

【0030】

【数35】

$$\hat{Q}$$

に実際に存在する要素のみとQの相当する要素とを比較する

【0031】

【数36】

$$\hat{Q}$$

に最も近いQ行列を見いだすことで、誤差を低減するアフィン構造および動きを見いだすことができる。そこで、欠落データの存在下に動きから構造を見いだす上で問題は、欠落要素を有するデータ行列に最も近い3階数行列を見いだすか否かによって決まる。

【0032】シャシュア (A. Shashua, 1992)、モーゼス (Y. Moses, 1993) およびベルフメウアラ (P. Belhumeur et al., 1996) が同様の式を用いて、無限遠の点光源に露光された3次元ランベルト構造によって形成される強度画像集合について説明している。その強度によって計算される光源の方向を3次元ベクトルhとする。その場合、 q_j を画像画素jの強度とする。最後に、この強度が、表面法線 n_j およびアルベド p_j を有する場合の3次元面状部分によって得られるものと仮定する。そうすると、強度が下記式によって得られる。 $q_j = h \cdot p_j n_j$ ただしこれは、正の値を提供するものである（すなわち、各表面放線が光源に向かっている）。これらの式から、hや p_j の絶対値を知ることはできないことから、 $p_j n_j$ が単に無制約3次元ベクトルである

と、一般性を失うことなく仮定することができる（すなわち、反射率が1未満であるという制約は、ほとんど制限なく、 h のスケールに対して影響がない）。次に、視点固定されているが、照明の強度および方向が可変である画像集合を考える。 j 番目の列が p_{jn} である場面行列を N とする。 i 番目の行が i 番目の画像についての h ベクトルである照明行列を H とする。さらに、 j 番目の列が各画像において j 番目の画素の強度を与えるデータ行列を

【0033】

【数37】

$$\hat{Q}$$

とする。そうすると、次の式が成り立つ。

【0034】

【数38】

$$\hat{Q} = HN + E$$

$$Q = \hat{Q} - E = HN$$

式中、 E はやはり、感知誤差の行列である。やはり Q は3階数を有する。やはり、我々の雑音性画像は、最大階数行列

【0035】

【数39】

$$\hat{Q}$$

を提供するものであり、やはり、SVDを用いて

【0036】

【数40】

$$\hat{Q}$$

から Q 、 H および N を推定して、

【0037】

【数41】

$$\hat{Q}$$

に最も近い3階数行列とそれの分解とを求めることができる。既報の研究で、シャシュア (A. Shashua, 1992) およびベルフメウアラ (P. Belhumeur et al., 1996) は、3つの画像のみを用いる比較的簡単であるが確実性の低い方法を用いて

【0038】

【数42】

$$\hat{Q}$$

を形成しており、その場合、

【0039】

【数43】

$$\hat{Q}$$

は3階数のみを持ち、 Q の最良の推定行列である。シャシュア (1992) は、行列 N を用いて、新たな画像が N の行の一次結合であるか否かを求めることで、その新たな画像が同じ場面によって形成され得るか否かを確認している。

【0040】この問題では、場面表面の法線が光源に向

かっていない場合に、

【0041】

【数44】

$$\hat{Q}$$

は欠陥要素を持つ。 $h \cdot p_{jn}$ が負の場合、その場面点は自己付影され、我々はゼロ強度とノイズを知覚し、負の強度は知覚しない。場面点が暗いことから、それが光源に向いていないとすることができるが、その方向性についてのそれ以上のデータは得られない。ベルフメウアラ (1996) は N を用いて、場面が作り出すことができ、その自己付影効果を可能とする画像集合について説明するより複雑な照明円錐形を求めているが、自己付影のない画像を用いて N をやはり構築しなければならない。しかしながら、自己付影を持たない多くの画像を得ることができれば、欠陥要素を持った

【0042】

【数45】

$$\hat{Q}$$

行列を構築することができる。やはり、対象の量を推定する上での重要な問題は、欠陥要素を有する行列に対する最適な3階数近似を見いだすということである。

【0043】トマシラ (1992) は、欠陥データを有する行列に対する最適な3階数近似を見いだす方法を提案している。その方法では最初に、欠陥データを持たない行列の直交する部分集合の位置を決定する。SVDを用いて、この小行列に対する最適な3階数近似を見いだす。次に、別の列（または行）の非欠陥要素に最も良く適合する小行列の列（または行）についての基底の一次結合を見いだすことで、その列（または行）の欠陥要素を満たす。このようにして、その解は、同時に一つの列（または行）に反復して拡張される。次に、最急降下最小化を用いて得られた解を最適化して、

【0044】

【数46】

$$\hat{Q}$$

行列に存在する全てのデータに最も近い3階数行列を見いだす。

【0045】過剰制限データを用いて欠陥データを満たしていることから、この解法はかなり発見的であるように思われ、トマシラ (1992) は、2つの動きシーケンスに対するその有効性を示している。しかしながら、欠陥行列要素を最初に満たす提案の方法はいくつかの欠点を有する可能性がある。第1に、欠陥要素を有する行列の最大の完全小行列を見いだす上での問題はNP完全問題であることから（最大クリークは容易にこの問題に帰することができる）、発見的手法を用いて開始点を見いださなければならない。第2に、データが対称に使用されない。データの小さい部分集合を用いて、最初の欠陥要素を計算する場合があります、それによってこれらの要素においてかなりの不正確さが生じる可能性がある。その不

正確さが別の欠陥要素の計算にどのように拡大するかについては明らかではない。第3に、 $m \times n$ 行列の場合、その方法はSVDのO (最大 (m, n)) 適用を必要とする。最後に、最急降下の最終適用により、その方法が全体的に最適解に収束するか否かは不明である。

【0046】シュムら (H. Shum et al. (1995)) も最近、欠落データ問題に関係する研究を行っている。彼らは、有意性を加重することができる3次元データに基づいて構造および動きを誘導する方法の問題を検討し、低階数行列をデータ行列に適合する上での加重最小自乗問題に到達している。この問題について彼らは、ワイベルグ (T. Wiberg, 1976) の報告に記載の知見による反復法の使用を提案している。この反復法は明らかに、この論文で検討されている欠陥データ問題にも適用することができる。この方法は、最急降下アプローチより良好な収束性を持つように思われ、我々の実験ではそれを用いている。この方法については、シュムらの報告 (H. Shum et al., 1995) に詳細に記載されている。しかしながら、我々の実験ではこの方法を用いることから、完全を期するため、その記載についてここで概要を説明する。ここではその方法は、最小自乗問題の双線形最適化として公式化し、次に一つの変数集合を一定としながら他の変数を最適化する方法を反復して、各最適化を線形とする方法として説明する。

【0047】詳細には、最初に、欠陥要素を持たない雑音性

【0048】

【数47】

$$\hat{Q}$$

行列から $Q = TP$ を計算したいとする。その式中、 Q は階数 r の $m \times n$ 行列であり、 T は $m \times r$ であり、 P は $r \times n$ である。 T および P の推定値を帰納的に改訂し、 T_i および P_i を各値についての我々の i 番目の推定値と称する。最初に、ある種の初期推定を用いて、 T_i を固定する。次に、最小自乗法の意味で $T_i P_i$ が

【0049】

【数48】

$$\hat{Q}$$

に最も近くなるような P_i を求めたい。これこそが、過剰制限線形予測問題であり、それは以下のようにして解決することができる。

【0050】

【数49】

$$P_i = T_i^+ \hat{Q}$$

式中、 T_i^+ は T_i の擬逆行列である。次に P_i を固定して、次の計算を行う。

【0051】

【数50】

$$T_{i+1} = \hat{Q} P_i^+$$

明らかに、これによって一連の推定値 Q 、 T および P が得られ、その場合の

【0052】

【数51】

$$\hat{Q} - Q$$

の大きさは増加しない。

【0053】わずかに変更を加えることで、

【0054】

【数52】

$$\hat{Q}$$

が欠陥要素を持つ場合に、その同じアルゴリズムを適用することができる。それを行うため、 $Q = TP$ を以下のように書き換えることができる。

【0055】

【数53】

$$\begin{pmatrix} Q_{11} \\ Q_{12} \\ \vdots \\ Q_{m1} \\ Q_{12} \\ Q_{22} \\ \vdots \\ Q_{m2} \\ \vdots \\ Q_{mn} \end{pmatrix} = \begin{bmatrix} T & 0 & 0 & 0 & 0 & 0 \\ 0 & T & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & T \end{bmatrix} \begin{pmatrix} P_{11} \\ P_{21} \\ \vdots \\ P_{r1} \\ P_{12} \\ \vdots \\ P_{r2} \\ \vdots \\ P_m \end{pmatrix}$$

式中、 Q および P はベクトルとして記載されており、 T は大きいブロック対角行列として記載されており、 T は n 回繰り返されている。やはり、 T を固定し、そのブロック対角行列の擬逆行列に

【0056】

【数54】

$$\hat{Q}$$

から構築されるベクトルを掛けることで、同じ方法を適用することができる。それによって、 P の新しい推定値が得られる。その反復法の第2の部分は、 T を簡単にベクトルとして書き、 P の同様のブロック対角行列を構築することで行うことができる。

【0057】そこで、

【0058】

【数55】

$$\hat{Q}$$

が欠陥要素を持つ場合に、この方法を容易に修正することができる。

【0059】

【数56】

Q

の欠陥要素に相当するベクトルおよびブロック対角行列の全ての行を削除し、その方法を前述のように行う。各ブロックの擬逆関数は別個に得ることができることから、そのブロック対角行列の擬逆行列を得ることはあまり経費を要しないことは留意すべき点である。

【0060】従って本発明の目的は、その問題の幾何学的構造に対する新たな洞察に基づいて、欠陥要素を持つ行列に対して低階数行列を適合させる新規な方法を提供することにある。我々の方法は、いくつかの望ましい性質を有する。ノイズがない場合には、我々の方法は正確な答えを正しく与えるか、あるいは悪化している場合はそれによって回答が得られないことを示すことができる。さらに、それは、完全に占有された行列の部分集合に依存するのではなく、全てのデータを利用できる。我々は実験的に、我々の方法がノイズの存在下に良好な結果を与えることを示している。これらの結果は、直接使用することができるか、あるいは反復方法の優れた開始点として役立ち得るものである。同時に我々は、良好な開始点なしで、反復方法で極小が追求される場合が多いことを示している。

【0061】従って、いくつかの欠陥要素を有するデータ行列に低階数行列を適合させることで、より信頼性が高く一般的な解を得ることができる。それにより、動きから構造のアルゴリズムで、一部の画像で見ることができない追跡場面点を用いることができ、場面が部分的にしか照明されていない多くの画像から場面の強度を構築することができる。

【0062】

【課題を解決するための手段】我々は、欠陥要素を有する行列に低階数行列を適合させる新規な方法を提供し、それを動きから構造問題と強度画像説明の問題の両方に適用する方法を示す。我々の方法は、異なっているが関係の深い関数型の最小化の問題に対する直接の解法に基づいたものである。

【0063】この方法は、いくつかの望ましい性質を持っている。第1に我々は、ノイズがない場合、我々の方法は正確な答えを正しく与えるか、あるいは悪化している場合はそれによって回答が得られないことを示すことができる。すなわち、適合行列が全ての非欠陥要素においてデータ行列と完全に一致する場合、我々が最小化を行う誤差関数型もゼロである。第2に、我々の方法は反復法ではない。第3に、それは、完全に占有された行列の部分集合に依存せずに、全てのデータを利用することができる。我々は実験的に、我々の方法がノイズの存在下に正確な結果を与えることを示している。これらの結果は、直接使用することができるか、あるいは反復解法の優れた開始点として役立ち得るものである。

【0064】従って、欠陥データを有する行列に低階数行列を適合させる新たなアルゴリズムが提供される。行列間の差の2乗の合計を最適化するための反復法に代えて、我々は、線形代数の標準的な手段、特にSVDを直接用いて最適化することができる新たなコスト関数を示すものである。その新たなコスト関数は発見的であるが、それが行列の差の2乗の合計と密接な関係を持つことを我々は示す。我々の説明は3つの部分からなる。第1に我々は、誤差のない場合のみを考えて、そのアルゴリズムの背景となる幾何学的直観について説明する。次に我々は、感知誤差を考慮した基本的アルゴリズムについて説明する。最後に我々は、今後の改良についての提言とともに、さらにそのアルゴリズムを強固なものとする多くの発見的改良について説明する。

【0065】さらに、不完全なデータを提供するビデオシーケンスから完全な場面構造を得る方法も提供される。

【0066】その方法の第1段階では、移動カメラを用いて固定場面の画像のシーケンスを得ることで、動きシーケンスから点位置からなる第1の行列を構築し、そのシーケンスから点の素性を確認・追跡し、その素性の座標を用いて、一部の素性が一部の画像に存在しない一部の欠陥要素を有する第1の行列を構築する。第2段階では、第1の行列から列のトリプルを選択し、その零空間を第2の行列とし、第2の行列の3つの最小成分を取ることで、近似解を得る。第3段階では、その3つの最小成分に反復アルゴリズムを適用して、第3の行列を構築し、推定を向上させる。最後に第4段階で、第3の行列を分解して、完全な場面構造を求める。

【0067】別の実施態様では、不完全データを提供するビデオシーケンスから完全な場面構造を得る方法が提供される。その方法の第1段階では、それぞれ異なった点光源を用いて固定場面の一連の画像を得て、付影または反射のために信頼性のないデータを与える各画像中の点を確認し、データの信頼性がない欠陥要素を有する強度値の第1の行列を構築することで、複数画像からの強度値からなる第1の行列を構築する。第2段階では、第1の行列から列のトリプルを選択し、その零空間を第2の行列とし、第2の行列の3つの最小成分を取ることで近似解を得る。第3段階では、その3つの最小成分に反復アルゴリズムを適用して、第3の行列を構築し、推定を向上させる。最後に第4段階で、第3の行列を分解して、完全な場面構造を求める。

【0068】本発明の別の態様は、不完全なデータを与えるビデオシーケンスから完全な場面構造を与える装置である。その装置は、固定場面の画像シーケンスを得るためのカメラと、そのシーケンスから点素性を確認、追跡する手段と、動きシーケンスから点位置からなる第1の行列を構築し、一部の素性が一部の画像で存在しない素性の座標を用いて一部の欠陥要素を有する第1の行列

を構築し、1) 第1の行列から列のトリプルを選択する段階、2) その零空間から第2の行列を得る段階、3) 第2の行列の3つの最小の成分を得る段階、4) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5) 第3の行列を分解して完全な場面構造を求める段階を有する近似解の構築を行う計算手段を有してなる。好ましい実施態様では、計算手段はパーソナルコンピュータである。

【0069】別の実施態様では、不完全データを提供するビデオシーケンスから完全な場面構造を得る装置が提供される。その装置は、固定場面の一連の画像を得るための複数の異なる点光源と、ならびに複数の画像からの強度値からなる第1の行列を構築し、付影または反射のために信頼性のないデータを提供する各画像中の点を確認し、データの信頼性がない欠陥データを有する強度値の第1の行列を構築し、1) 第1の行列から列トリプルを選択する段階、2) その零空間から第2の行列を得る段階、3) 第2の行列の3つの最小の成分を得る段階、4) その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5) 第3の行列を分解して完全な場面構造を求める段階を有する近似解の構築を行う計算手段を有してなる。好ましい実施態様では、計算手段はパーソナルコンピュータである。

【0070】

【発明の実施の形態】本発明の装置および方法についてのこれらおよび他の特徴、態様および利点に関しては、以下の説明、添付の特許請求の範囲、ならびに添付の図面を参照することで理解を深めることができる。

【0071】図1には、実際の動きシーケンスの1フレームを示してある。

【0072】図2(a), (b)には、2つの比較法による、5つの解法の平均誤差を示してある。各点は、100回の試験で平均した、データと3階数適合との間の誤差の2乗の合計を示してある。平均誤差を、ある点が入っているフレームの平均部分fに対してプロットしてある。適合と比較するデータが相対的に少ないことから、誤差を(1-f)で割ることで、誤差低減の合計を補正する。図2(a)のプロットの数値目盛のプロットには、4つの解法を示してある。図2(b)では、誤差がかなり少ない均等目盛法でプロットしている。比較のため、両方のプロットでGT-ITを示している。

【0073】図3は、最適な利用可能解に収斂する反復法の回数の割合を示している。

【0074】図4は、100回の平均での、欠陥データを有する強度画像に3階数行列に適合させる実験で得られた平均残留誤差を示したものである。

【0075】図5は、反復法が強度画像についての最適利用可能解に収斂する頻度を示したものである。

【0076】図6(a), (b)のそれぞれは、図2

(a), (b)に示したような実際の動きシーケンスについての平均誤差を示したものである。

【0077】図7は、最適な利用可能解に収斂する反復法の回数の割合を示している。

【0078】図8は、本発明の方法を実行する装置を示したものである。

【0079】図9は、本発明の別の方法を実行する別の装置を示したものである。

【0080】我々の問題の幾何学的直観を得るには、

【0081】

【数57】

\hat{Q}

の各列を、m次元空間における点の座標を与えるものと見ることが役立つ。この考え方で、この行列が欠陥要素を持たない場合に、トマシラ(C. Tomasi et al., 1992)またはウルマンら(S. Ullman et al., 1991)がSVDを用いて、

【0082】

【数58】

\hat{Q}

におけるm次元点に「最も近い」3次元線形副空間Lを見いだしている。Lは、それがこれらの点に対する距離の2乗の合計を最小とするという点で、これらの点に最も近い。同時にSVDは、

【0083】

【数59】

\hat{Q}

で相当する点に最も近いLにおける各点を見いだすことで、Qの列を構築する。

【0084】

【数60】

\hat{Q}

が欠陥要素を有する場合、これらの要素を満たして、モデルとデータの間の適合を改善することができる。欠陥要素を有する

【0085】

【数61】

\hat{Q}

の列は、一部未知の座標を有する点である。すなわちそれは、m次元空間におけるアフィン副空間である(そして、この空間の軸に対して平行である)。この場合の我々の問題は、互いの距離の2乗の合計によって求められる、これらアフィン空間に最も近い3次元線形副空間Lを見いだすことである。Qおよび

【0086】

【数62】

\hat{Q}

の欠陥要素は、Qの各列を

【0087】

【数63】

19
 \hat{Q}

の相当する列によって表されるアフィン空間に最も近いLでの点に割り付けることで見いだすことができ、Qのこの列の欠陥要素は、Lに最も近いこのアフィン副空間でその点を取ることで満たすことができる。

【0088】ここで我々は、感知誤差がない場合（例：
【0089】
【数64】

\hat{Q}

が、欠陥要素以外はQと同一である場合）に適切なアル
10 ゴリズムを構築する方法を示し、次に、そのアルゴリズム
に変更を加えて、誤差を考慮する方法について説明す
る。ノイズがない場合、Lは

【0090】
【数65】

\hat{Q}

の列によって記述されるアフィン空間全てと交差する。
交差しないこれらのアフィン副空間の3つを A_1 、 A_2 、
 A_3 と称する。空間が交差しないというこの仮定は実際
には意味がないことに留意すべきである。アフィン空間
は座標軸に対して平行であることから、2つの空間は、
全ての行についてそれらが同じ値を持つか、あるいは一
方が欠陥値を持つ場合にのみ交差する。そのようなトリ
プルは、上述のアルゴリズムでは容易に捨てられる。L
はそれらのアフィン空間それぞれと共通の原点以外の特
有の点を有することから、Lはこれら3つの空間におけ
る点によって囲まれた空間内にあるに違いない。従っ
て、

【0091】
【数66】

\hat{Q}

の列の各トリプルは、Lを R^m の何らかの線形副空間に
制限する。この制限の意義は、これら列における欠陥要
素の数によって決まる。列が欠陥要素を持たない場合、
それらは正確にLを規定するものである。別法として、
列が多く欠陥要素を有する場合、それらが生じるアフ
ィン空間は、高次元のものであり、非常に高次元の空間
を囲むと考えられる。この場合、Lがその空間にあると
いう制限は弱い制限である。

【0092】トリプルから生じる線形空間の交差部を取
ることで、
【0093】
【数67】

\hat{Q}

の列の多くの異なるトリプルに由来する制限を組み合わ
せることができる。具体的には、

【0094】
【数68】

\hat{Q}

の列に相当するアフィン空間の何らかのトリプル集合を

20

Tとし、アフィン空間 A_1 、 A_2 、 A_3 における点によっ
て囲まれる線形空間をスパン (span) (A_1 、 A_2 、
 A_3) とすると、次のようになる。

$$\begin{aligned} & \text{【0095】} \\ & \text{【数69】} \\ & L \subseteq \bigcap_{(i,j,k) \in T} \text{span}(A_i, A_j, A_k) \quad (2) \end{aligned}$$

この場合部分集合を得ているのであって、同等の関係で
はないことに留意する。それはすなわち、

$$\begin{aligned} & \text{【0096】} \\ & \text{【数70】} \\ & \bigcap_{(i,j,k) \in T} \text{span}(A_i, A_j, A_k) \end{aligned}$$

が3次元である場合に、Lは正確にわかるが、その交差
部がそれより高次元である場合、正確にはわからないと
いうことである。さらに、Qの階数が不明確である場合
には、

$$\begin{aligned} & \text{【0097】} \\ & \text{【数71】} \\ & \bigcap_{(i,j,k) \in T} \text{span}(A_i, A_j, A_k) \end{aligned}$$

はその階数のより上の界を与えるものである。実際に
は、我々のシミュレーションは、誤差のない場合に、か
なりの量の欠陥データについて、

【0098】
【数72】

\hat{Q}

の列の可能なトリプルの小さい部分集合のみのサンプ
リングを行うことで、Qを個別的かつ正確に決定すること
ができる。

【0099】実際のアルゴリズムでは、感知誤差の可能
性を考慮しなければならない。誤差があると、各列に関
連するアフィン空間を横切ること、我々のデータと完全
に一致し得る線形副空間がないことから、

$$\begin{aligned} & \text{【0100】} \\ & \text{【数73】} \\ & \bigcap_{(i,j,k) \in T} \text{span}(A_i, A_j, A_k) \end{aligned}$$

は直ちに空となる。そこでその場合、この交差部にでき
るだけ近くなる線形副空間を計算する。

【0101】最初に、
【0102】
【数74】

\hat{Q}

の列の1トリプルを選択する。この時点では、単にこれ
らが無作為に選択している。列のi番目のトリプルによ
って囲まれる線形空間を S_i とする。これらの空間の交
差部を取る簡便な方法は、その補空間のスパンの余集合
を取るものである。 S_i の零空間を N_i と称する。次に、
誤差のない場合に、

【0103】

【数75】

$$L \subseteq \text{零空間}(\text{span}_{i=1}^l N_i)$$

を得る。下記の行列の零空間を取ることで、

【0104】

【数76】

$$\text{零空間}(\text{span}_{i=1}^l N_i)$$

を計算することができる。

【0105】 $N = [N_1 N_2 \dots N_l]$

感知誤差の存在下に、行列 N は最大階数を有するのが普通であり、その零空間は空である。その場合、その零空間に近い3次元線形空間、すなわち少量のノイズを考慮することで、零空間に最も容易に形成される空間を見いだしたい。それは、 N のSVDを取り、その4つの最下位成分を見いだすことが行われる。それにより、 N によって囲まれる空間に対して最も直交に近い3次元線形副空間として L の推定を行うことができる。

【0106】 N のSVDの3つの最下位成分は、 L を囲む3つの正規直交列として L を表すものである。次に、それらを用いて Q を制限する。

【0107】

【数77】

$$\hat{Q}$$

の各列について、欠落していない要素と最も適合する L を表す列の一次結合を見いだす。具体的には、 L を3つの正規直交性列によって表し、

【0108】

【数78】

$$\hat{Q}$$

の列を

【0109】

【数79】

$$\hat{Q}$$

で表し、

【0110】

【数80】

$$\hat{Q}$$

の非欠陥要素の指数を p とし、

【0111】

【数81】

$$\hat{Q}$$

の非欠陥要素のみを取ることで得られる列を

【0112】

【数82】

$$\hat{Q}$$

とし、同様に p の指数を有する行からなる L の小行列を L^p とする。そうすると、次のように計算される。

【0113】

【数83】

$$Q_i \approx L((L^p)^+ + \hat{Q}^p)$$

式中、 $(L^p)^+$ は L^p の擬逆行列である。

【0114】ここで、このアルゴリズムを実行する上で重要ないくつかの詳細について説明する。第1に、3つのアフィン空間によって囲まれる空間の零空間を計算する方法について説明する。3つのアフィン空間 A_i 、 A_j 、 A_k を生じる

10 【0115】

【数84】

$$\hat{Q}$$

の列トリプルを C_i 、 C_j 、 C_k とする。 C_i 、 C_j または C_k のいずれの行が一つの要素を欠落している場合、その行はスパン(A_i 、 A_j 、 A_k)の零空間の全ての要素について零値を持つことに注目する。従って、その零空間を計算するには、完全に占有されて、零空間の残りの行に全て零が入った $[C_i C_j C_k]$ の小行列の零空間を見いだす必要があるのみである。

20 【0116】第2に、スパン(A_i 、 A_j 、 A_k)の零空間は、数値的不安定さのために、 L に対する正確な制限ではない。簡単な例として、 A_i 、 A_j 、 A_k が、その原点とともに、ほとんど何らかの2次元平面 P にある3つの点であると仮定する。そうすると、そのスパンは3次元であり、感知誤差がない場合には、 L を正確に規定すると考えられる。しかしながら実際には、 P を含む3次元線形副空間は、3点全ての近くにあり、ノイズが加わると、そのスパンは、 P を含むかなり任意の3次元線形副空間である。従って、 L がそのスパンにあると間違っ

30

て予想されるものと考えられる。しかしながらこの状況は、容易に検出される。我々は、完全に占有された $[C_i C_j C_k]$ の小行列に対してSVDを適用することで、零空間を計算する。この小行列が最大階数でないかあるいはほとんど最大階数でない場合、それに由来する制限が信頼性のないものであることがわかっている。

40

【0117】第3に、我々は N を計算するのに用いた列トリプルの数に1という閾値を設定しているが、これら3つのそれぞれがどれだけの数の列を与えるかは事前にはわからない。従って、必要な N の列数に対して最大値も設定し、1個のトリプルをチェックしていなかったとしても、その値に達した時に、列トリプルの形成を停止する。

【0118】第4に、式2は部分集合の関係に過ぎないことから、得られた列トリプルが、正確に L を決定できるだけそれを制限しない可能性が常にある。 L は、その3つの最小特異値に基づいて、 N の3つの最下位成分を取ることで計算されることに再度留意する。 N の第4の最小特異値も非常に小さい場合、 L についての計算が信頼性のないものであることがわかっている。その場合は単に、アルゴリズムによって答えが得られなかったと報

50

告される。ただし、1についてのさらに大きい値または列トリプルの別の無作為選抜によって、より多くのデータが得られる可能性を期待して、計算を再度行う場合も考えられる。

【0119】第5に、

【0120】

【数85】

\hat{Q}

の全ての列および行を求めることができるだけのデータをNが直接含まない可能性がある。例えば、選択される列の全てのトリプルがNのある行において1以上の欠陥要素を有している場合、その行は全て零からなる。その場合、計算している零空間におけるその行の値についてのデータを集めたことになる。さらに、前述のように、 (L^p) が最大階数を持たず、その擬逆行列を得ることができないことから、Lによって Q_i を推定することはできない可能性がある。それらの状況のいずれにおいても、単純に、データを持っているQの最大亜矩形(subrectangle)を計算し、その亜矩形を拡大して、完全矩形の最小自乗計算を行う。

【0121】図8について説明すると、参照符号800によって表される本発明の装置を図示してある。図8は、不完全なデータを提供するビデオシーケンスから完全な場面構造を形成する装置800を図示してあり、図中、装置800は固定場面810の画像シーケンスを得るためのカメラ820を有している。さらに、シーケンスを介して点素性を確認・追跡する確認・追跡手段830が提供されている。好ましくは、確認・追跡手段830は、トマシら(C. Tomasi and T. Kanabe, "Detection and Tracking of Point Features," Carnegie-Mellon Tech. Report CMU-CS-91-132, 1991)およびコックスら

(I. Cox and S. Hingorani, "An Efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and its Evaluation for the Purpose of Visual Tracking," IEEE Trans on PAMI, vol. 18, no. 2, pp. 138-150, Feb. 1996)の報告に記載の方法などの当業界で公知のアルゴリズムである。

【0122】装置800にはさらに、動きシーケンスから点位置からなる第1の行列を構築し、一部の素性が一部の画像で存在しない素性の座標を用いて一部の欠陥要素を有する第1の行列を構築し、1)第1の行列から列のトリプルを選択する段階、2)その零空間から第2の行列を得る段階、3)第2の行列の3つの最小の成分を得る段階、4)その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5)第3の行列を分解して完全な場面構造を求める段階を有してなる近似解の構築を行う計算手段が与えられている。計算手段840は好ましくは、パーソナルコンピュータである。

【0123】図9について説明すると、参照符号900

で表される本発明の別の装置が図示してある。図9は、不完全なデータを提供するビデオシーケンスから完全な場面構造を形成する装置900を図示してある。この装置900は、固定場面810の一連の画像を得るための独立の複数の点光源910を有してなる。

【0124】装置900にはさらに、複数の画像からの強度値からなる第1の行列を構築し、付影または反射のために信頼性のないデータを提供する各画像中の点を確認し、データの信頼性がない欠陥データを有する強度値の第1の行列を構築し、1)第1の行列から列トリプルを選択する段階、2)その零空間から第2の行列を得る段階、3)第2の行列の3つの最小の成分を得る段階、4)その3つの最小成分に反復アルゴリズムを適用して第3の行列の構築および推定の向上を行う段階、5)第3の行列を分解して完全な場面構造を求める段階を有してなる近似解の構築を行う計算手段がある。計算手段は好ましくは、パーソナルコンピュータである。

【0125】

【実施例】この項では、本発明の方法の正確さを調べるための実験について説明する。本発明の方法および従来の方法は、データ行列に存在する要素と新たな3階数行列の相当する要素の間の差の2乗の合計を最小化するという同じ問題処方を検討するものである。従って、方法によって階数3行列を見いだして、その誤差測定を低減することにどの程度奏功するかに基づいて、それらの方法を評価する。独立のガウス雑音がデータの信頼性を低下させる場合に、この誤差測定は最適であり、欠陥要素の値について事前にデータが得られないことは留意すべき点である。

【0126】次のように標識した7つの異なる方法で実験を行う。すなわち、それ自体で本発明の方法を使用する「NEW」。反復開放の開始点として使用する「NEW-IT」。これを、無作為開始点による反復開放と比較する。良好な開始点を持つことの重要性を明らかにするため、その反復法を1つの無作為開始点(RAN1)、3つの異なる無作為開始点(RAN3)、あるいは5つの異なる無作為開始点(RAN5)で行う場合に認められる最良の解を得る。我々はこれらの結果を(GT)グラウンドトルースの推定値、および(GT-IT)反復法の開始点として(GT)での解を用いることで認められる全体的に最適な解(誤差が最も小さいもの)の推定値とを比較する。シミュレーションしたデータを用いる時には、グラウンドトルースが使用できる。実際の動きシーケンスでは、(GT)についての全てのデータに基づいた解を用いる。

【0127】これらの実験では、多くの特定の閾値を設定しなければならない。100回の反復で、あるいはそれが10桁の精度で収束するまでその反復法を行う。本発明のアルゴリズムでは、元の行列の10倍の列または行を零空間(N)が占有するまで、列の3つを選択する

(列の3つが1000個という上限があるが、実際にはそれに到達することはほとんどない)。列の3つそれぞれに対する零空間を計算する場合、零および0.1未満の特異値を得る。NのSVDの3つの最下位成分を取ることでLを計算する。第4の最小特異値が0.001未満である場合、その行列に対する明瞭な3次元零空間がないことから、結果に信頼性がないことを求めている。

【0128】2つの設定で合成データを得る。第1に、合成動きシーケンスで点を得る。各シーケンスで、立方体内で無作為に20点を選択する。次に、対象をz軸について90°回転させ(面内回転)、その平面で無作為に選択される軸について90°回転させながら、均等の間隔で20のフレームを得る。これによって、動きアルゴリズムからの構造が共通に適用されるシーケンスの種類の簡単な例を提供するものである。次に、各点がシーケンスのある部分に閉塞されると仮定する。この部分は、均等分布から各点について無作為に選ばれる。閉塞長さの予想値を変化させて、我々の実験における欠陥要素の数を変える。各点についての閉塞フレームは、シーケンスの始め、シーケンスの終わり、または始めと終わりの間で分割されるかのいずれかに無作為に割り付けられる。これら3つの条件のそれぞれが、等しい確率で起こる。次に、要素の最大範囲の、25%の標準偏差により、ガウス雑音を加える。500×500画素の画像の場合、それは1.25画素の標準偏差に相当すると考えられる。

【0129】次に、無作為強度画像を得る。最初に、1000個の点を有する場面を得て、そのうちの全ての場面の表面法線を、正のz成分を有する単位半球での均等な分布で選択する。0~1の均等な分布から各点の反射率を選択する。次に、単位強度の光源を選択し、場面表面法線と同じ分布から引き出される方向を選択する。各光源をその場面と組み合わせ、-1~1の範囲の強度で強度画像を得る。次に、標準偏差、0.1でガウス雑音を加えた。最後に、零未満の強度を欠陥値と見なした。

【0130】最後に、ソーニーらの報告(H. Sawhney et al. (1990) "Description and Reconstruction from Image Trajectories and Rotational Motion" Int. Conf. on Comp. Vis: 494-498)に記載の実際の動きシーケンスから得られたデータについて、我々のアルゴリズムを行った。このシーケンスは、8個の画像にわたって追跡された40の点を含んでいた。このシーケンスでは、無作為動きシーケンスについて説明したものと全く同じ人工的閉塞が生じた。次に、隣接するフレームに存在する多数の点の中心の移動を測定し、その移動を除去することで、移動の効果を補償した。次に、その行列のスケールを決定することで、合成動きシーケンスの場合と同じ次数とすることで、同じパラメータを全て適切に使用できるようにした。これにより、全ての方法で、感知誤差および移動を求める上での不正確さによる誤差を含ん

だ同じ行列で作業が行われるようにした。

【0131】図2、3、4、5、6および7では、結果をグラフ表示している。第1の注目点は、我々の方法を反復解法の開始点として用いた場合に最良の結果が得られているという点である。動き実験においてこの方法は、無作為開始点を用いて反復解法を1回、3回または5回繰り返した場合よりかなり良好な結果を与える。詳細には、閉塞されるデータが40%または50%で、問題がより困難になるに連れて、我々の方法の成績の低下は、複数の無作為開始点を用いる反復解法の場合よりかなり遅い。強度画像への線形表面の適合においては、その問題は無作為開始点を用いた場合にはかなり困難であると考えられるものであるにも関わらず、我々の方法はほぼ完全な結果を与える。

【0132】第2に、反復による向上を行わずに、我々の基本的方法によって良好な結果が得られる場合があることがわかる。動きの場合、点の20~30%が閉塞されている場合、我々の方法は、グラウンドトゥースにおける誤差の約2倍の誤差で解を与える。強度実験の場合、我々の方法は、グラウンドトゥースの場合よりわずかに30%高い平均誤差で解を与える。これらの数字は、我々の方法が、反復による向上を行わずに、良好な近似解を与え得ることを示している。しかしながら、動きにおける閉塞量が増えると、我々の解の質はかなり低下し得ることは留意すべき点である。

【0133】第3に、かなりの量のデータ欠陥がある場合、動き問題は非常に困難なものとなり得ることがわかる。我々の方法であっても、データの半分が欠落している場合の90%よりわずかに高い最良解を与えるのみである。動き問題の困難さは、欠陥データの数だけでなく、特定の閉塞パターンや動きの大きさおよびフレームの数によっても決まるものである。しかしながら、我々の実験は妥当な流れに基づいたものであると考えられ、動きがさらに少なく閉塞がさらに大きいというさらに困難な動き状況は興味深いものと考えられる。従って我々のシミュレーションは、予備的なものではあるが、非常に高い確率で最適解が望まれる場合、対象とする現実の問題は、既存のアルゴリズムに対して困難な問題を提示するものであることを示している。

【0134】欠陥要素を有するデータ行列の線形適合が、最近のいくつかのアルゴリズムの中心の問題である。我々は、反復解法の開始点として特に有用な、この問題を解決する新規な方法を提供するものである。さらに我々の実験では、一部の現実の設定では、この問題は容易ではなく、反復解法が良好な開始点を持たずに不正確な結果を与える場合が非常に多いことが示されている。

【0135】以上、本発明の好ましい実施態様および実施例と考えられるものについて説明したが、当然のことながら、形式や詳細において各種の修正および変更を本

発明の精神から逸脱することなく容易に行うことができることは明らかである。従って本発明は、以上説明した通りの形のものに限定されるものではなく、添付の特許請求の範囲内にあると考えられる全ての変更を含む構成のものであるべきである。

【0136】

【発明の効果】本発明は以上説明したように構成されているので、以下に記載するような効果を奏する。

【0137】欠陥要素を持つ行列に対して低階数行列を適合させることができノイズがない場合には、正確な答えを正しく与えるか、あるいは悪化している場合はそれによって回答が得られないことを示すことができる。

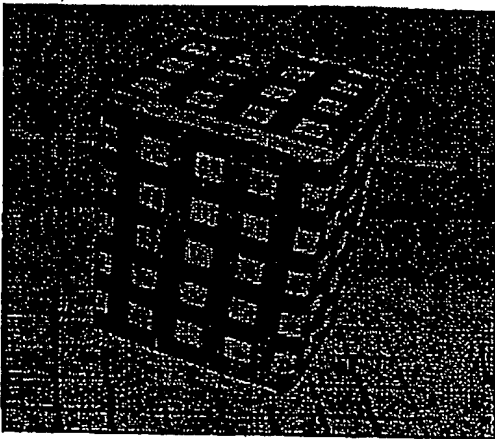
【0138】従って、いくつかの欠陥要素を有するデータ行列に低階数行列を適合させることで、より信頼性が高く一般的な解を得ることができ、動きから構造のアルゴリズムで、一部の画像で見ることができない追跡場面点を用いることができ、場面が部分的にしか照明されていない多くの画像から場面の強度を構築することができる効果がある。

【図面の簡単な説明】

【図1】実際の動きシーケンスの1フレームを示してある。

【図2】(a), (b)は、2つの比較法による、5つの解法の平均誤差を示してある。各点は、100回の試験で平均した、データと3階数適合との間の誤差の2乗の合計を示してある。平均誤差を、ある点が入っているフレームの平均部分fに対してプロットしてある。適合と比較するデータが相対的に少ないことから、誤差を $(1-f)$ で割ることで、誤差低減の合計を補正する。

【図1】



(a) のプロットのプロットには、4つの解法を示してある。(b)では、誤差がかなり少ない均等目盛法でプロットしている。比較のため、両方のプロットでGT-ITを示している。

【図3】最適な利用可能解に収束する反復法の回数の割合を示している。

【図4】100回の平均での、欠陥データを有する強度画像に3階数行列に適合させる実験で得られた平均残留誤差を示したものである。

10 【図5】反復法が強度画像についての最適利用可能解に収束する頻度を示したものである。

【図6】(a), (b)のそれぞれは図2(a), (b)に示したような実際の動きシーケンスについての平均誤差を示したものである。

【図7】最適な利用可能解に収束する反復法の回数の割合を示している。

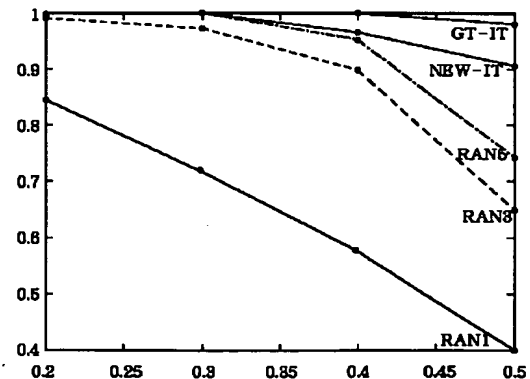
【図8】本発明の方法を実行する装置を示したものである。

20 【図9】本発明の別の方法を実行する別の装置を示したものである。

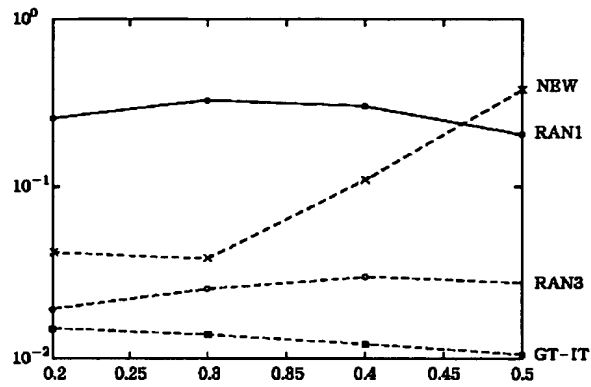
【符号の説明】

800	装置
810	固定場面
820	カメラ
830	確認・追跡手段
840	計算手段
900	装置
910	点光源

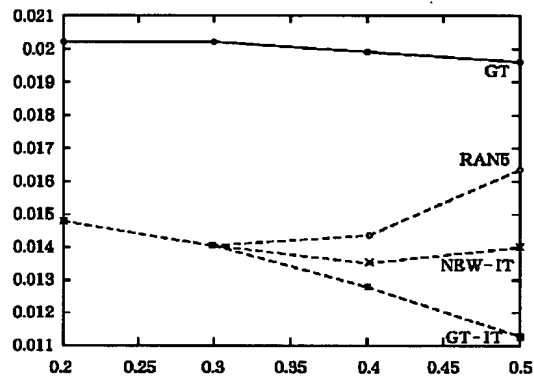
【図3】



【図2】

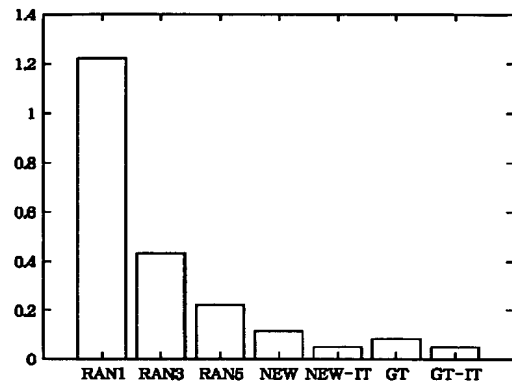


(a)

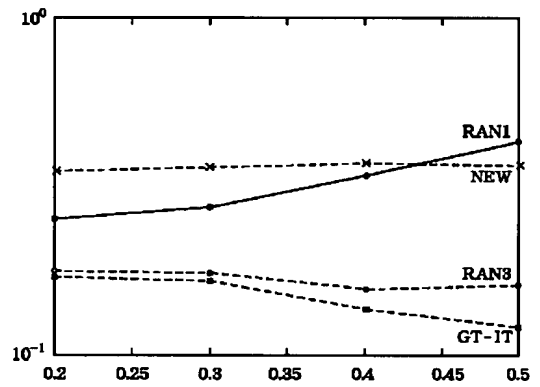


(b)

【図4】

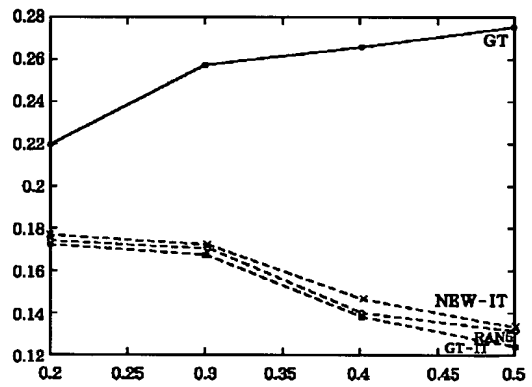
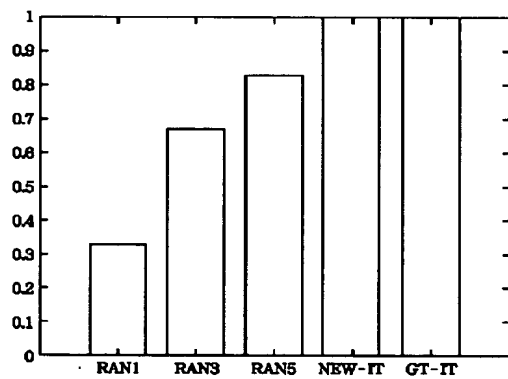


【図6】



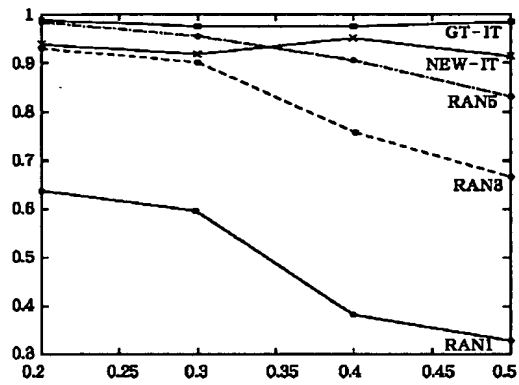
(a)

【図5】

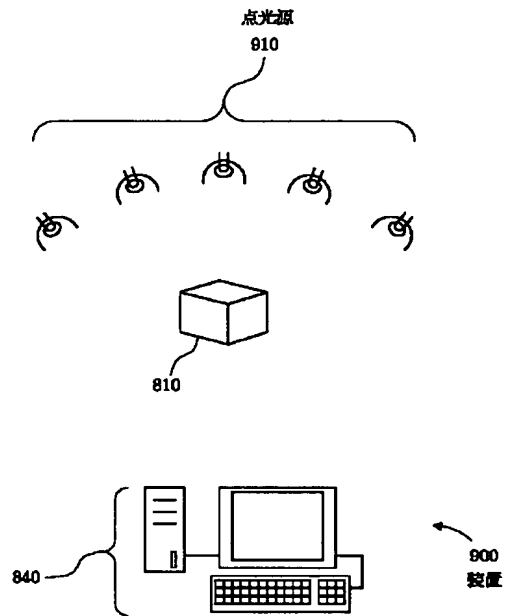


(b)

【図7】



【図9】



【図8】

